Journal of the Royal Statistical Society
SERIES A Statistics in Society

J. R. Statist. Soc. A (2011)
174, *Part* 3, *pp.* 547–567

Discussion on the meeting on 'Resource allocation models'

Irini Moustaki (*London School of Economics and Political Science*)

The paper by Galbraith and Stone provides a critique of the regression tools that are used in selecting and interpreting the resource allocation formulae that the Department of Health is using for funding primary care trusts. It is also a critique of the reports produced and the justification that they provided for modelling and fitting.

The aim of the funding formula as it is described in the paper is to provide equal access to people at equal risk, making two assumptions that are considered to be unrealistic in practice by the authors and other researchers.

The paper focuses on the validity of the statistical tools that are used, mainly p-values, t-tests, R^2-values, the regression specification error test statistic and bootstrapping techniques as a method of cross-validation. Test statistics in all the reports have been treated at face value and the authors show that often highly statistically significant parameters turn out to explain a very small proportion of the total variance. In addition the use of bootstrapping for cross-validation when both the model is misspecified and the data are fixed will produce erroneous results. The paper addresses an extremely interesting and important area. The critique of the statistical tools is of course very important and it applies to many other applications as well. I can identify three areas of concern.

First, the model that was adopted by the Department of Health does not have a strong theoretical justification as becomes apparent from the discussion and related papers. Variability across regions needs to be taken into account and modelled by region level socio-economic variables. A better justified model from health economic theory will not have received such severe criticism even in the case of statistical evidence that is not strong. Often with social science data only a moderate proportion of variability is explained by the explanatory variables but researchers are still happy to work with models as long as they rely on some social theory.

Second, variables such as 'needs' and 'supply' variables do not seem to be measured accurately and even some components of those variables are missing from the model. Better measures are required and if those are not easy to measure one might need to consider involving constructs such as latent variables in the model. This procedure will separate the measurement process from the rest and statistical tools such as factor analysis and item response theory can help to detect problems with specific indicators. The paper does not provide any more information on the need and supply variables. Some more discussion on the measurements used would be useful for the reader.

Third, Galbraith and Stone make an effort to signal the failure of a misspecified model. There is no surprise in the existing findings. The critique is very welcome and well founded. However, I think that in many instances they ought to discuss alternative ways of overcoming the existing problems, especially if we assume that those findings can be also observed in a situation where the model is justified by some theoretical framework. Readers should use the findings in a broader perspective than the Department of Health's funding formulae and more guidelines are needed in this case. Would the parametric bootstrapping have been justified if the model was a better approximation to reality? Is the regression specification error test statistic a good way of judging the lack of model misspecification? There are many published papers that deal with those questions but readers would also need some references and guidelines in the current paper.

Finally, Galbraith and Stone do not provide ways of improving the formulae and the regression model. I understand that this might be out of the scope of the paper but some ideas would be worth discussing and investigating.

Finally, I thank the authors for choosing to work on such an important area of research with immediate implications for our society. I feel that their work ought to be better acknowledged by the relevant government departments and the researchers who are involved.

I am more than happy to propose the vote of thanks.

© 2011 Royal Statistical Society
0964–1998/11/174547

Mike Heiser (*Local Government Association, London*) (©Local Government Association)
I am grateful for the opportunity to make a few remarks. I should begin by saying that they are on a personal basis and should not be taken as a statement of policy by the Local Government Association.

In seconding the vote of thanks, I shall concentrate on the local government grant system as opposed to the health system. I would first like to talk about the legal framework within which it operates and the considerations that are likely to influence Ministers and the officials who advise them. Secondly I would like to look in more detail at the development and the workings of the current formula grant distribution system. Thirdly I would like to look into the future and to suggest some ways in which the system might change.

I should say that I am not going to enter into an extensive discussion of statistical formulae, for which, in this august company, apologies are due. However, I hope that I might shed some light on how statisticians and statistical formulae can find themselves constrained by non-statistical considerations.

The Local Government Finance Act 1988 requires that the Secretary of State, every year, issues a report specifying the basis on which he will distribute revenue support grant and business rates. That basis is set out in the *Local Government Finance Report* which must be approved by Parliament annually. It also must be, in a word, judge proof. Although in the past the courts have proved extremely reluctant to second-guess the basis on which the Secretary of State chooses to distribute, as long as, in the words of Lord Justice Bridge in a landmark 1990 case, it can be shown that there is no bad faith, improper motive or manifest absurdity (Regina *versus* Secretary of State for the Environment, ex p. Hammersmith and Fulham LBC [1991] UKHL 3 (October 4th, 1990) at `http://www.bailii.org/uk/cases/UKHL/1991/3.html`), no Secretary of State wants to push that too far. In practice, therefore, the basis that is referred to in the Act is a mixture of objective formulae, with a statistical basis, constrained by judgement.

As has been set out, the basis on which the formula grant is currently distributed is popularly known as the 'four-block model'. That in summary is as follows:

(a) an element for relative needs which is intended to reflect the relative costs of providing comparable services between different authorities;
(b) an element for relative resources which takes account of the different capacity of different areas to raise income from council tax due to the differing mix of properties; it is a negative amount as it represents assumed income for authorities;
(c) a central allocation which is the same per head for all authorities delivering the same services;
(d) a floor damping block to give every authority a minimum grant increase, or maximum grant decrease; grant changes to other authorities are scaled back to pay for this.

Thresholds per head are used to split measures of absolute need and absolute resources into the relative needs, relative resources and central allocation.

The internal structure of the absolute needs measures (confusingly called by the Department for Communities and Local Government (DCLG) relative needs formulae) is not substantially different from the system of standard spending assessments which applied from 1990–1991 to 2002–2003 or from the formula spending shares which applied from 2003–2004 to 2005–2006; both, in general, constitute a per-client amount with uplifts for deprivation and geographical characteristics, such as sparsity and higher labour costs.

However, both formula spending shares and standard spending assessments were expressed in cash terms—as the amount of spending which the government was willing to support through grants. The formula grant was derived as follows:

$$\text{formula grant (before damping)} = \text{formula spending share} - \text{tax base} \times \text{assumed national council tax}.$$

Fig. 1 sets this out graphically in a simplified form, assuming that there is only one tier of authorities and that their needs and resources vary uniformly.

The current way of determining the formula grant has departed from any system of cash spending assessments and no longer contains an explicit figure for assumed national council tax. However, the change is largely presentational, as the needs and resources blocks and the central allocation can be recast in the old form: Fig. 2.

Clearly, as has been pointed out, if you scale relative needs or the central allocation to a predetermined total rather than 1 you change the relativities in this diagram. In 2010–2011 the central allocation is scaled up by 9.1%; in 2011–2012 it is scaled down by 36.6%. In terms of Fig. 2 you are increasing or decreasing the rectangle AHDG.

But that is only one of several areas where the statistical formulae are constrained by judgement. They include

Downloaded from https://academic.oup.com/jrsssa/article/174/3/547/7077760 by guest on 05 November 2025

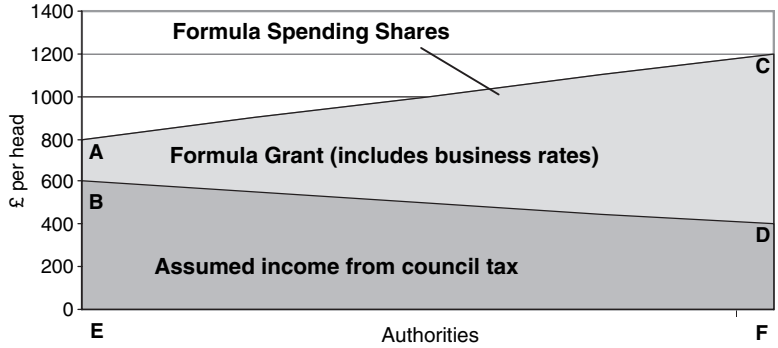

Fig. 1. Formula spending shares (ACFE) reflect the differences in need per head; assumed income from council tax (BDFE) reflects the differences in resources; formula grant including redistributed business rates (before damping) (ACDB) is the difference between the two

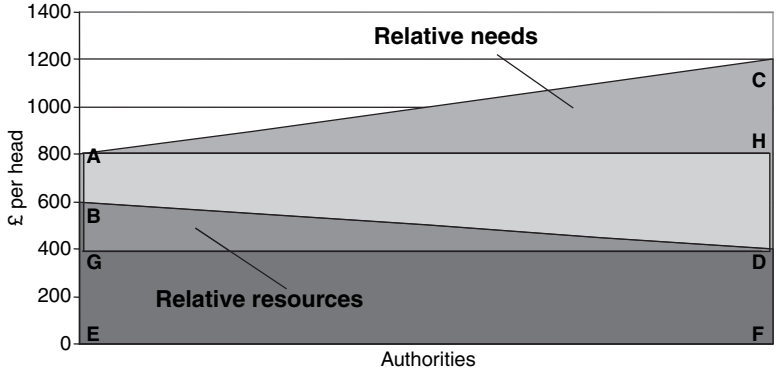

Fig. 2. Same simplified model under the post 2006–2007 formula grant system: the formula grant before damping remains the same (ACDB); it is now made up of relative needs (ACH) plus the central allocation (AHDG) less relative resources (BDG); the central allocation (AHDG) is the difference between the amount below threshold for the relative needs block (AHFE) and that for the relative resources block (GDFE)

(a) the total size of the needs and resources block and the central allocation and thus their weight relative to one another,
(b) the tier shares in the relative resources block,
(c) the weight of the different service blocks relative to each other within relative needs formulae and
(d) the 'floors' and scaling factor that are used for damping.

And yet DCLG statisticians spend inordinate time in developing models to fit those parts of the system which do have some statistical basis. For example, in developing a fire risk index in 2010 they looked at about 90 possible variables, using a mixture of principal component analysis and linear regression.

All the drawbacks of the current system were displayed during the formula grant consultation of July 2010. For example, a change affecting how concessionary fares are funded and administered led to about 30 exemplifications being included in the consultation, all of which purported to show how the formula grant would have been different in 2010–2011 if the particular change being exemplified had been in place. Because of the workings of the four-block model and because of the amount of judgement that was used in setting tiers splits and assumed damping levels for the different exemplifications it was impossible for authorities to link the change to any one cause, and thus very difficult for them to respond meaningfully to the consultation. The particular change that was ultimately proposed by the DCLG within the 2011–2012 settlement satisfied very few and, as the Local Government Association said in its response to the consultation, seems likely to lead to a whole new set of winners and losers.

Downloaded from https://academic.oup.com/jrsssa/article/174/3/547/7077760 by guest on 05 November 2025

Many think that the forthcoming local government resource review provides an opportunity to reform the formula grant system, and, to judge by their recent pronouncements, that includes DCLG Ministers. Currently, most of what purports to be formula grant is actually redistributed business rates. Councils might get to retain more of their business rates locally rather than paying the whole amount into a pool. However, there could still be an equalization mechanism so councils which collect much in business rates would pay into a smaller pool and those in areas which have a lower business rate tax base would receive a payment from the pool. There is thus a continuing case for an overall needs assessment. But many would hope that it was simplified from the current relative needs formula and put back into a form which is more comprehensible to key officers and members.

The vote of thanks was passed by acclamation.

Mervyn Stone (*Ruislip*)
Before 'damping', the grant of £G_a to authority a with population n_a can be written

$$G_a = \left(n_a g_a / \sum_a n_a g_a\right) B_1 - \left(n_a d_a / \sum_a n_a d_a\right) B_2 + p_a B_3$$

where

(a) B_1, B_2 and B_3 are prefixed billions with B_2 less than both B_1 and B_3,
(b) $g_a = \Sigma_t\{f_{at} - \min_a(f_{at})\}$ where f_{at} is a socio-econometric proxy formula for *per capita need* in need tier t,
(c) $d_a = \Sigma_t\{c_{at} - \min_a(c_{at})\}$ where c_{at} is the *per capita council tax base* for resource tier t and
(d) $\{p_a\}$ are proportions whose specification by the Department for Communities and Local Government aims, but fails, to achieve a uniformly equal *per capita* distribution of B_3 over England.

Wokingham is the minimizing authority in formula (b) for the 'upper' tier $t = 1$. The 'Wokingham effect' on authority a is the change in the coefficients of B_1 and B_3 when Wokingham's f_{W1} is imagined raised to the value of Richmond upon Thames's f_{R1}—the next lowest value in tier 1. The messy algebra for this can be graphically evaluated for all 456 authorities, but that would do little more than throw one or two lumens onto a stage that is already well lit by this bold and entertaining critique of resource equalization hubris. It is now over 3 years since Alex Gibson told me about his intra-ocular impacting discovery—to which (I find) the nigh-universal reaction of ordinary folk is either concussed silence or explosive laughter. Those with a foot in public administration tend to be more circumspect, whatever their political persuasion. It may therefore be unrealistic to expect a *mea culpa* from the government. Indeed, while writing this, I heard some Ministerial talk of 'merging some floors'.

Wynand P. M. M. van de Ven (*Erasmus University Rotterdam*)

A new challenge for the resource allocation formula in the English National Health Service: how to prevent risk selection by the new general practitioner consortia
The resource allocation formula in the English National Health Service is already very complex, as Galbraith and Stone, and Gibson and Asthana indicate. The proposed reforms (see for example Her Majesty's Government (2010) and Timmins (2010)) make it even more complicated:

(a) the primary care trusts are to be abolished;
(b) about 75% of the National Health Service's budget will be allocated to general practitioner (GP) consortia;
(c) consumers will have a free choice of GP.

A new challenge is: *how to prevent risk selection by the new GP consortia*.

If the consumers have a choice between different GPs and different GP consortia, the allocation formula must be calculated at the individual consumer level, rather than at the small area level (see for example Bevan and van de Ven (2010)). If one individual moves from one GP consortium to another, this person's budget must follow the consumer. Therefore, the appropriateness of the allocation formula must be judged at the *individual consumer* level, rather than at the *small area* level. This further complicates the allocation formula. Undercompensation and overcompensation of certain risk groups provide the GP consortia with incentives for risk selection, which may have several unfavourable effects.

First, GP consortia then have a disincentive to respond to the preferences of high risk consumers. For

Downloaded from https://academic.oup.com/jrsssa/article/174/3/547/7077760 by guest on 05 November 2025

Table 1. Dutch risk equalization formula 2011

In 2011 the following risk adjusters are used (see for example van de Ven and Ellis (2000), van de Ven *et al.* (2004) and Prinsze and Van Vliet (2007):

age and gender (38 groups);
pharmacy-based cost groups, a morbidity measure based on information about chronic conditions deduced from outpatient prescribed drugs (23 groups);
diagnostic cost groups, a morbidity measure based on information about the diseases diagnosed during previous hospitalizations (13 groups);
source of income and age ($4 \times 4 = 16$ groups);
socio-economic status and age, 12 groups based on interaction of

(a) 4 groups based on average income and address,
  (i) 1 group, if more than 15 people per address,
  (ii) 3 groups based on average income per address, if not more than 15 people per address,
(b) 3 age groups;

region (10 clusters of postal codes)

example, GP consortia with a good reputation for chronic care would attract many unprofitable patients and would be the victim of their own success (e.g. Beaulieu *et al*. (2006)). Therefore, GP consortia may choose not to contract with hospitals that have the best reputation for treating chronic illnesses. Second, efficient GP consortia, which do not engage in risk selection, may lose market share to inefficient risk selecting GP consortia, resulting in a welfare loss to society.

In the Netherlands we have 20 years of experience with resource allocation at the individual level. The equalization payment that a Dutch insurer receives for an individual in a certain subgroup is based on the predicted national average health expenses *per capita* in this subgroup. According to a recent evaluation this Dutch equalization formula (Table 1), which is one of the most sophisticated in the world, severely undercompensates several groups of high risk individuals, resulting in substantial incentives for selection (Table 2). Therefore, the proposed health reforms provide England with a new challenge: how do they prevent risk selection by the new GP consortia?

Ian Hunt (*Edinburgh*)

Galbraith and Stone's general claim is that data mining is an abuse of regression. By this they mean that stepwise regression selections are invalid—especially if the steps follow 'classical statistics' like R^2, Student's t and Fisher's F. I agree that the regression modelling of Morris *et al.* (2007) is data mining done badly, but I believe that recent econometric research shows that data mining is not bad *per se*.

Lovell (1983) warned, as do Galbraith and Stone, that data mining leads inevitably to incorrect models and inferences. However, in pages 171–264 of volume 7 of the *Journal of Economic Methodology* (2000) econometricians argue that data mining (and searching over model spaces) is widespread, useful and sometimes even essential. I think that the main problem is that econometricians often pretend they are not data mining when they know that they are ('giggling uncomfortably when caught' (McCloskey (1985), page 201)): using classical statistics outside a classical setting should be justified explicitly.

The *ad hoc* data mining of Morris *et al*. (2007) might not work properly, but evidence from modern econometrics methodology is more encouraging. In fact, stepping 'backwards and forwards and sideways, before the final selection of any model' (as Galbraith and Stone pejoratively put it) can be done well—especially with modern computing power. For example Hendry (Hendry and Granger, 2005) uses a semi-automated approach to remove parameters from general 'encompassing models' via rules that are based on the very classical statistics which Galbraith and Stone admonish. More generally, Spanos (1999, 2000) argues that a researcher can be licensed to data-mine when her focus is on inferential error and 'statistical adequacy' of candidate models.

Alongside their anti-data-mining message, Galbraith and Stone suggest that statistical analysis without a 'scientific basis' or 'convincing scientific argument' is wrong headed. Do they mean that statisticians should not search for theories and models, or analyse data outside the context of theory? When Maurice

Downloaded from https://academic.oup.com/jrsssa/article/174/3/547/7077760 by guest on 05 November 2025

Table 2. Average predictable financial result (profit or loss) per adult in year t for selected groups after applying the Dutch risk equalization formula for 2007, given community-rated premiums (excluding the *ex post* cost-based compensations and fixed hospital costs)†

Subgroups based on information of year $t-3$ or earlier	*Size of subgrouping (%)*	*Predicatable financial result (2007 €) in year t*
Self-reported health status (year $t-3$)		
Self-reported health status fair or poor	21.2	−541‡
10% worst score physical functioning (short form 36)	10	−1140‡
Auditive impairment (Organisation for Economic Co-operation and Development score)	4.8	−308
Visual impairment (Organisation for Economic Co-operation and Development score)	7.3	−222
Mobility impairment (Organisation for Economic Co-operation and Development score)	14.9	−653‡
10% worst score mental health (short form 36)	10.0	−297‡
Stroke, brain haemorrhage or infarction	2.6	−943‡
Myocardial infarction	3.3	−789‡
Some type of malignant cancer	4.8	−689‡
High blood pressure	15.2	−342‡
Vascular constriction (stomach, legs)	4.1	−947‡
Asthma, chronic obstructive pulmonary disease	8.1	−460‡
Dizziness when falling down	6.2	−615‡
Intestinal obstructions (>3 months)	4.3	−484‡
Urinary incontinence	7.7	−500‡
Osteoarthritis (hip, knees)	16.3	−401‡
Chronic joint inflammation	6.1	−725‡
3–6 self-reported conditions	22.3	−333‡
7 or more self-reported conditions	2.9	−1461‡
Prior healthcare utilization (year $t-3$)		
Use of prescribed medicine in previous 14 days	48.2	−220‡
Contact with medical specialist in previous year	39.8	−317‡
Hospitalization in previous year	7.5	−1034‡
Contact with physiotherapist in previous year	18.5	−315‡
Home healthcare in previous year	2.2	−1152‡
Top 25% with highest expenses (year $t-7$, $t-3$)		
4 out of 5 years	4, 5	−304‡
All 5 years	8, 2	−1757‡
Hospitalization (year $t-7$, $t-3$)		
2 out of 5 years	4, 7	−728‡
3 out of 5 years	1, 1	−2030‡
4 out of 5 years	0, 3	−5933‡
All 5 years	0, 1	−6453‡

†Source: Stam and van de Ven (2007, 2008).
‡Statistically significantly different from 0 ($p<0.05$).

Kendall was told that theory must come before statistics he responded '[this] seems to me the negation of the scientific spirit' (Kendall (1953), page 34). If modern statistics 'is the art of discovery' (Hand (2008), page 114), an extensive exploration of covariates and models (*à la* data mining) may play a critical creative role.

Inadequate statistical modelling of the National Health Service funding allocations is nothing to giggle about. Galbraith and Stone conclude that the ultimate (and unaddressed) issues are the real world significance of estimated parameters and the fragility of inferences in light of alternatives and benchmarks. Long time econometric campaigners McCloskey (1985, 2008) and Leamer (1983) would agree.

Downloaded from https://academic.oup.com/jrsssa/article/174/3/547/7077760 by guest on 05 November 2025

Nigel Williams (*Epsom*)

Complexity in local government allocations can divert attention from the regressiveness at the centre of the council tax. Before the marginal effects of benefits and reliefs, the largest band H property, with no upper limit on its valuation, is charged only three times the smallest in band A in the same local authority, valued from zero upwards (Department for Communities and Local Government, 2008).

The National Health Service allocation formula is more complex still. It lists some 50 variables on which the calculation is based (Department of Health, 2008). In describing just one, these remarks hope to underline the cumulative level of complexity.

The housing domain from the English indices of deprivation, which were produced in 2000, is used to predict mental health need. It is a composite statistic combining homeless households in temporary accommodation, overcrowding and poor private sector housing. Overcrowding data came from the latest available census, 1991 (Department of Environment, Transport and Regions, 2000). An equivalent, subtly different statistic has been produced twice since. Housing domain no longer exists in its own right, forcing a choice between revisiting the selection process for variables and using an increasingly outdated data set. The housing domain was not produced as a direct measure of mental health. Rather, it and the many companion indices serve to show relative levels of deprivation between small geographical areas. Grant-making bodies typically look at ranks. They determine a threshold level at which they consider there is scope for improvement and then treat areas at that rank or worse as eligible. Grants may be made where the threshold is exceeded, with less emphasis on by how much.

The indices of deprivation make only limited claims to be to scale. Some are actual rates, but this was never true of housing. The separate indicators, once combined into a single domain, could not readily be treated as having units: hence the emphasis on dealing with ranks. Only when combining different indices is some sort of scale imposed. The ranks are fitted to a common exponential distribution with the most deprived according to each index occupying the long tail (Noble *et al.*, 2004). That transformation represents the index compilers' view of the best way to fit deprivation to a scale. They are concerned that severe deprivation in any domain should not be masked by a process of averaging. The resulting index serves better as an identifier of exceptions than as a scaled measure of need.

The 2011–2012 National Health Service allocation formula omits the housing domain in its calculations (Fillingham, 2010).

Darrin L. Baines (*University of Warwick, Coventry*)

Deriving ought from is: is the approach to constructing National Health Service funding formula terminally flawed?

Concerns over the construction of the National Health Service resources allocation formula are not new. Since the introduction of the Resource Allocation Working Party formula in the late 1970s, the limitations of basing financial allocations on statistically derived weights have been well known. For instance, the Resource Allocation Working Party was criticized for failing to measure need and for failing to take account of the problems of inner cities (Ashton, 1978; McCarthy, 1986). Most National Health Service funding models introduced in the last 30 years have been criticized because of the statistical methods that they employed. During the 1990s formulae for deriving prescribing budgets were criticized for the deficiencies of

(a) using data available rather than data that measure the real drivers of costs,
(b) measuring correlates of spending and not variations in need and
(c) basing weights on historic resource use patterns without fully understanding their desirability (Baines and Parry, 2000).

Given these persistent and damaging criticisms, we must ask why does the health service still use such a flawed method?

In *The Tyranny of Numbers*, David Boyle suggested that the modern preoccupation with numbers has created a 'pervasive blindness' in our perception of the world, so we no longer see the non-quantitative character of life (Boyle, 2001). In the context of the National Health Service, we could argue that the availability of high powered software packages and massive data sets has encouraged the belief that we can obtain simple answers from complicated correlations. However, when undertaking such work, the enormous number of data points and high levels of multicollinearity make the construction of 'appropriate' statistical models almost impossible. Often, models with the right signs and significances (which tell the required story) are chosen, with little public acknowledgement of the major compromises that have been

Downloaded from https://academic.oup.com/jrsssa/article/174/3/547/7077760 by guest on 05 November 2025

made in their construction. This occurs for three main reasons. First, policy makers incorrectly believe that they must employ statistical methods for their formula to be scientifically valid. Next, researchers have failed to build more sophisticated models of health service operation than simple regressions. Finally, we have wrongly rejected 'historic cost funding' as inappropriate, when, in fact, the past has much to tell us about how to fund future services. In conclusion, we should consider whether the problem lies deeper than misapplied statistics. In using descriptive data to construct prescriptive weights, researchers are falling foul of the 'naturalistic fallacy' in the sense that they are inappropriately trying to derive 'ought from is' (Baines and Parry, 2000). In response, perhaps statistical methods should be put aside, while we find more sensible approaches to determining the funding needs of local populations.

Woody Caan (*Anglia Ruskin University, Cambridge*)
In various ways, I have been involved in National Health Service commissioning since *Working for Patients* in 1989 (Department of Health, 1989). This included work on 'Shadow purchasing Project 26', general practitioner (GP) fund-holding, training fledgling primary care groups and then primary care trust commissioners, as well as my old-fashioned health authority duties.

It is urgent that new approaches to allocating funds are developed, as indicated by Galbraith and Stone. The imminent transfer of £80 billion to untried GP consortia makes better systems even more essential. If, after this transfer the new budget 'follows the consumer' there is a real risk to public health. During my past research around GP behaviour towards drug users (Rozewicz *et al*., 1992) or gypsy travellers (Caan, 2006) it became clear that a large majority of these clinicians were reluctant to address the needs of these 'high risk consumers', as Wynand has mentioned earlier. Patients do not even need to threaten any *identifiable* extra costs, to evoke risk avoidant behaviour in general practice, as observed towards people with learning disabilities (Caan, 2009).

However, my greatest concern about future commissioning (involving perhaps 300 or more consortia) is that the Health and Social Care Bill does not require any coherent geographical distribution across England: a group of medical pals can club together as and where they like. Data on the health needs and care experiences of their 'consumers' will not correspond to any current National Health Service or local authority information areas. In the larger health authority system we struggled to employ even one qualified statistician per commissioner (let alone other numerate professions like health geographers and economists). In 2011 most primary care trusts do not have a qualified statistician. Can the Royal Statistical Society help the GPs by educating a *new cadre* of health statisticians who can understand rapidly evolving funding allocations and then relate them to whole new patient data systems? Help!

Greg Phillpotts (*Dorking*)
The two papers we have heard today show very clearly that the world of resource allocation has become extremely complex. As Sir Kenneth Stowe, then Permanent Secretary of the Department for Health and Social Security said in the early 1980s, 'If this confuses you, that shows you are thinking clearly'.

With these complex systems, as shown by these papers, it is clearly vital that the allocation process can be demonstrated to be operating correctly in terms of its objectives, as well as in terms of comparisons between authorities or areas and for the same area over time. There is obviously a need quickly to identify allocation systems that are not right and hence which result in some authorities receiving 'too much' and others 'too little' money as assessed by the method's own indicators of need. It is also important to do this validation promptly so that there can be some time to correct matters before the next development comes along. The signs are that there will be another National Health Service system shortly, so it is important to move ahead with work on such a diagnostic method promptly.

The kind of demonstration that I have in mind is not the same thing as the debate about the theoretical correctness of methods, although that is also important. What is called for is more of an indication or perhaps a series of tests that can be carried out on the calculated allocations and the needs and supply data. An example given in the presentations today is the undue influence of values for an individual authority or a small number of authorities; no doubt there are a range of others. I would look to graphical methods, which would help to make the results accessible at a variety of levels of government.

Joan Davis (*The Community Voice, London*)
As a non-statistician I am privileged to be present and permitted to speak on a topic that has concerned me for many years. The Community Voice is an umbrella organization that is dedicated to provision of

Downloaded from https://academic.oup.com/jrsssa/article/174/3/547/7077760 by guest on 05 November 2025

good National Health Service (NHS) services locally. Our 45 member organizations—residents associations, patient groups, branches of the three major political parties—come from north-west London and south-west Hertfordshire, and together represent many thousands of ordinary people.

We have long been concerned about the NHS funding formulae. The 2003–2006 formula had 19 proxy variables, leading to huge differences in allocation per head of population—twice as much for every person in inner city Tower Hamlets as for their counterparts in more rural Wokingham, in Surrey. That worried us. We do not begrudge support for Tower Hamlets people, but we believe that their health inequalities stem from factors that are not addressed by throwing money at health services—factors such as poverty, unemployment, poor diet, housing and education. In contrast there are issues relating to provision of services that can and should be addressed by additional funding, such as aged buildings and local staffing difficulties.

However, there is no way that ordinary people can assess complex statistical formulae, whether they are fair or not—only the Royal Statistical Society has both the national authority and the breadth of expertise to undertake such work.

Since then the formula has changed but—as Stone and Galbraith confirmed—the new formula is equally complex, so ordinary people are still equally powerless to help themselves.

I therefore beg the Society and its members to use their unique right to speak on statistical issues. I ask them to champion ordinary people by initiating serious scrutiny of the NHS funding formulae, which dictate how vast sums of public money are distributed across our country.

In hoping that the Society will rise to my challenge, I reflect that this is the only organization with the independence, high ethics and stature to match that task.

Jon Spain (*London*)

The abuse of regression in National Health Service allocation formulae

Thank you for allowing me to comment on the interesting paper by Galbraith and Stone. I am an actuary, speaking personally and not on behalf of my employer. As Joan Davis has just said, the Royal Statistical Society (RSS) are ideally placed to provide advice and I believe that the RSS should do just that. Let me make just two comments.

First, outside the ranks of accountants, actuaries and statisticians, many people are staggeringly innumerate. Sadly, many of them are almost proud of it. This lack of numeracy can also be found within the civil service and among government Ministers. So the message needs to be as simple as possible but no simpler.

Secondly, if the RSS wish to have any influence, they must refrain from being critical because that just will not work. Instead, the RSS should offer diplomatic assistance, which may well be accepted: good luck and thank you!

J. R. Cuthbert and M. Cuthbert (*University of Strathclyde, Glasgow*)

The paper by Gibson and Asthana graphically illustrates the limitations of the current four-block distribution system for revenue support to local authorities in England. It is clear from the analysis in the paper that, far from providing a distribution which is fairly and demonstrably based on objective measures of need, the methodology that is used is impenetrable and arbitrary in its effects.

The purpose of this comment is not so much to deal with Gibson and Asthana's paper itself, but to point out important implications which their paper has for proposals which have been put forward for reform of the Barnett formula, i.e. the mechanism that is currently used to allocate funding to the devolved governments and assemblies of the UK. In December 2009, the Independent Commission on Funding and Finance for Wales, chaired by Gerald Holtham, suggested a replacement for the Barnett formula which took as its starting point the current (2010–2011) resource allocations to regions and areas within England, Scotland and Wales—including, importantly, local government allocations, as well as allocations for health, schools, etc. These allocations were used as the dependent variable in a regression to assign weights to different needs indicators, and it was these needs indicators which formed the basis of the proposed resource allocation to the devolved administrations. Holtham himself described this process, of deriving indicators from current policy, as a process of 'revealed preference'. The limitations which Gibson and Asthana's paper reveals in the four-block distribution method for English local authority expenditure indicate just how unsatisfactory the current English local authority distribution is as an input to Holtham's proposed methodology, and this therefore calls into serious question the applicability of the approach that was suggested by Holtham.

Downloaded from https://academic.oup.com/jrsssa/article/174/3/547/7077760 by guest on 05 November 2025

The 'Comprehensive spending review' that was published in October 2010 indicated that the UK Government was intending to consider with the Welsh Assembly the implications of adopting the Holtham methodology. The paper by Gibson and Asthana is a timely reminder that at least one of the key building blocks in this process is not fit for purpose.

C. Sanderson (*London School of Hygiene and Tropical Medicine*)
Firstly, I should declare an interest as a long-standing member of a committee that advises the Department of Health about resource allocation methods and related research.

I shall only comment on the alternative approach to resource allocation that was sketched out in the discussion. I interpret this as follows:

(a) define a number of care groups within which members benefit from similar packages of care;
(b) for each local population, count or estimate the number n in each care group;
(c) estimate the average cost c of each package;
(d) for each local population sum nc over all groups.

Health officials have been aware of the conceptual shortcomings of adjusted utilization for many years and in 2000 commissioned work on the feasibility of this 'bottom up' alternative. Briefly the conclusions were as follows: apart from constant returns to scale it had obvious theoretical attractions; however, it would be enormously demanding of both data and judgement; data would improve, so feasibility should be reviewed periodically; but such an approach could not be comprehensive and would have to be combined with other methods.

Let us dig deeper. If need for *healthcare* is related to potential benefit and risk, it is not sufficient to define care groups in terms of pathology alone. Studying need for hip replacement, Willcock (1979) found 46 cases of osteoarthrosis of the hip in general practitioner records. However, 22 were not sufficiently severe for surgery to be indicated, 11 were not sufficiently fit, four had already had surgery, five were otherwise ineligible and one did not want it. Only three were actually 'in need'. Sanderson *et al.* (l997) estimated that, in a population of 250 000, around 11 500 men would have moderate or severe urinary symptoms, 7600 with indications for prostatectomy; but only about 600 would probably or definitely want surgery. Estimates were highly sensitive to controversial threshold values.

Suppose nonetheless that we could agree a set of care groups, each with indications for treatment (and thus, implicitly, intergroup priorities). General practitioners could be asked to report the numbers n in their practices. However, ascertainment may depend, inequitably, on practice organization and care seeking behaviour. This also introduces an incentive to over-report. Alternatively we might estimate ns by using practice population data with models derived from survey data. Apart from the challenges of all this modelling, survey data are vulnerable to inequitable bias through sample selection, response and self-reporting of disease status, and to imprecision for uncommon but high cost categories: and what of patient acceptability?

Then what kinds of costs should be used?; national averages?; averages for care consistent with guidelines?; delivered efficiently?; data from where, updated how often?

In their interim approach, Galbraith and Stone are unduly sceptical. Weather forecasting has its faults but who would base policy on the assumption that every day will bring annual average amounts of sun and rain?

Adjusted utilization has recognized shortcomings but remains a practical approach with plausible results.

D. J. Bartholomew (*Sudbury*)
The allocation of very large sums of money according to perceived need is so common in the public service that it is surprising that statisticians have not made a bigger contribution. My own, rather marginal, involvement dates back to the days of the rate support grant. The object then was to devise a formula which could both be justified statistically and be credible to those who made the allocation. The method that was chosen was multiple regression which was as inappropriate then as now. *Ad hoc* adjustments had to be made because real 'need' could hardly have a negative coefficient! The rationale behind the method was to suppose that previous allocations had been made by using a linear function of a set of indicators of 'need'. The coefficients of that linear function were then estimated by least squares regression. The fit was generally poor which certainly suggested that the allocations had not been made by using such a formula!

It is true that the linear assumption may be reasonable as the famous paper by Cook (1976) shows. He deduced that the allocations made by the old University Grants Committee were made *as if* they were

Downloaded from https://academic.oup.com/jrsssa/article/174/3/547/7077760 by guest on 05 November 2025

based on a regression formula used by the Province of Ontario. A related approach, allowing a degree of feedback, was suggested by Bennett and Tan (1979).

There is a rather obvious alternative method which ought to have occurred to statisticians. Why not attempt to construct a single index of 'need' from the various indicators of 'need'? All one would then have to assume was that previous allocations had been made according to 'need'. If this relationship appeared to be linear, or at least monotonic, it could be estimated and used to make new allocations. The statistical apparatus for doing this has been available for many years.

One must sympathize with official statisticians given the task of advising on major allocations in a limited time. It is easy to be critical, as today's authors have rightly been. But Galbraith and Stone have argued before that a more direct approach could be adopted and with this suggestion I heartily agree. Fortunately, we do not have to look very far for such a method, though it would be foolish to expect that its application would be as straightforward as my rather simplistic account may have suggested.

Downloaded from https://academic.oup.com/jrsssa/article/174/3/547/7077760 by guest on 05 November 2025

Stephen Senn (*University of Glasgow*)

I found these interesting papers both encouraging and depressing: encouraging because they show the continued contribution, through learned papers such as these, that this great Society of ours can make to matters of public policy but depressing, because they reminded me of issues that were debated many years ago. To strike a personal note, one of the first papers I wrote, when working for the Tunbridge Wells Health District, was to criticize the system that was proposed in the 1970s for allocating resources to the National Health Service (Senn and Shaw, 1978). My comments are mainly applicable to Galbraith and Stone but also have some relevance to Gibson and Asthana.

A trivial point that statisticians need to make clear to others is that if regression-based allocation formulae are used then districts are gainers or losers depending on whether their current revenue is above or below the conditional fitted mean. This mean will reflect current patterns and thus what 'is' but not necessarily what 'ought to be'. Thus a regression model can indicate points that are unusual given an assumed model but this is not automatically an indication that the points need to be changed!

If I have a criticism it is that the authors were rather more convincing in saying what not to do than in proposing suggestions about what one should do. Yet, as Bayesians are wont to remind us, to refuse to make a decision is to make one and not necessarily a good one. To quote a contemporary of the (now) famous 'Minister of Tunbridge Wells', 'I call two physicians; they differ in opinion. I am not to lie down, and die between them: I must do something' (Samuel Johnson). For both papers, we have a very clear idea of one particular course of action that the authors would *not* take but this leaves a rather wide choice of alternatives and I hope in due course that we shall hear more on this score. Any public servant might justifiably argue that, if the way they were supposed to find an adequate policy was to keep on changing it until eventually they found one that no statistician would criticize, it might take rather a long time to come up with a solution. Do the authors have some positive advice for the unenlightened?

The following contributions were received in writing after the meeting.

Hugh Gravelle (*University of York*), **Stephen Morris** (*University College London*), **Nigel Rice** (*University of York*), **Peter C. Smith** (*Imperial College London*) **and Matthew Sutton** (*University of Manchester*)

We write as lead researchers on several of the projects that are referred to by Galbraith and Stone. These have led to major improvements in matching resources to local needs and have informed formulae to allocate National Health Service funds in England since 1995.

Both the tone and the content of the paper are regrettable. It is replete with errors of fact, misinterpretations of the criticized reports and misunderstandings of the policy process. We reject most of the criticisms of the statistical methods. It is impossible to provide a full rebuttal within the imposed word limit and we shall provide a detailed response in due course.

The two 'fundamental assumptions' set out in the paper are straw men and assumption A is a distortion of statements made in the reports cited. The purpose of the modelling process is to explain the average costs of healthcare arising from current patterns of behaviour, given characteristics of providers ('supply factors') and citizens and the small areas in which they live ('needs factors'). There is no suggestion that the observed patterns are optimal. In particular, the potential for the existence of 'unmet need' has been examined exhaustively and was one of the motivations for a 'health inequalities' adjustment applied to various formulae.

Both needs and supply factors reflect characteristics, drawn from epidemiological, economic and other literature, that are known to be associated with the need for medical care or that affect utilization of health

services independently of need. The model selection process is used to develop parsimonious models, which are comprehensible and acceptable to National Health Service policy makers, that can be used for policy purposes without any material loss of predictive accuracy. When calculating target allocations by using the predictive model, the effect of supply factors is removed by equating them to the national average level of supply. The formulae therefore predict healthcare costs in a small area, given its needs characteristics, but assuming that it has national average clinical practice and national average supply. Assumption B is an unnecessary obfuscation.

The paper's policy conclusion that current knowledge cannot justify a departure from uniform *per capita* allocations is absurd. Current National Health Service funding formulae are based on exhaustive analysis and detailed technical scrutiny, and are in our view fit for purpose. There is always scope for further improvement in data and methods following informed and constructive discussion. Unfortunately, this captious paper offers little contribution to this important debate.

Paul Hewson and Joyce Halliday (*University of Plymouth*)

Gibson and Asthana raise issues about the way in which the funding formula for local government has been developed. We have examined the allied formulae for fire and rescue and police funding which reinforce these concerns.

Our first issue is transparency. Suitable reports, explaining the derivation of the formula, were unobtainable. As Gibson and Asthana report, it is possible to find extensive, cross-linked and contradictory spreadsheets, and exemplar calculations. Yet, here our research suggests a more communicable formula can be substituted:

$$
\begin{aligned}
Y_{\text{Fire}} = \text{Scale}\, X_{\text{Area}} \Bigg\{ & 4.6998 + \underbrace{0.7788 \frac{X_{\text{Coast}}}{\text{Pop}_{2009}}}_{\text{Coastline}} + 0.233 \Bigg(\underbrace{\frac{X_{\text{Ch}} - 0.1929}{0.0682} + \frac{X_{\text{Ho}} - 0.7214}{0.0406} + \frac{X_{\text{R}} - 0.2571}{0.0657}}_{\text{Risk index} \ldots} \\
& + \underbrace{\frac{X_{\text{Ab}} - 21.6666}{1.3578} + \frac{X_{\text{Av}} - 2.2578}{0.1389} + \frac{X_{\text{Ac}} - 0.0211}{0.0111}}_{\ldots \text{ risk index}} \Bigg) + 30669.492300 \underbrace{\frac{X_{\text{Comah}}}{\text{Pop2004}}}_{\text{COMAH}} \\
& + \frac{865.2528}{\text{Pop2009}} \underbrace{X_{\text{BuildRisk}}}_{\text{Property and social risk}} + \frac{0.4132}{\text{Pop2007}} \Bigg(\underbrace{5X_{\text{Primary2004}} + \frac{X_{\text{ACORN risk}}}{X_{\text{ACORN}}} \text{Pop2004} + X_{\text{Pop2004} \geqslant 65}}_{\text{Community fire safety}} \Bigg) \Bigg\} \\
& - 4.4940954747 \qquad (1)
\end{aligned}
$$

where the data are defined in Table 3 and the key constants in Table 4 apply.

Hiding calculations and terms such as these behind complex spreadsheets restricts understanding and constrains challenge. Users need to know why terms are included and what each formula is supposed to model. Here, we believe the formula was based on a prediction of demand expressed by the number of calls, yet confirmation was never easy. Principal components and automatic model choice algorithms were apparently used in variable screening. One of the 'deprivation' indicators appears to enter the funding formula with the wrong sign, effectively removing money from deprived areas, and issues with multicollinearity and outliers draw the validity of the formula further into question.

Secondly, we are uncomfortable with the match between measured needs and funding requirements. The fire and rescue formula should be one of the simplest subcomponents of the four-block formula, yet there are considerable variations in funding allocated *per capita* and this is even more pronounced per call. It is difficult to see any conceptual justification for this variation.

Thirdly, the basis for measuring need does not always relate clearly to service burden. The fire and rescue need appears to be funded on 'all calls' received (Fig. 3). This contains a perverse incentive: rewarding those who fail to deal with false alarms. Conversely, no allowance is made for attending fatal road accidents, despite availability of the data. If the intention is to fund certain aspects of public service and not others this is a valid political decision. Transparent formulae would ensure that it is evident and capable of debate.

Downloaded from https://academic.oup.com/jrsssa/article/174/3/547/7077760 by guest on 05 November 2025

Table 3. Definition of data in formula (1)

Parameter	*Definition*
X_{Coast}	Coastline
X_{Ch}	Average number of children of income support or income-based claimants as a proportion of residents under 18 years old
X_{Ho}	Average proportion of households other than those containing married or cohabiting couple who have no children and no others
X_{R}	Average proportion of people in rented accommodation
X_{Ab}	Average number of absences from primary school
X_{Av}	Average number of rooms per household resident
X_{Ac}	Proportion of household residents in 'A classification of residential neighbourhood' 50 and 53
X_{COMAH}	Number of control of accidents and major hazards sites
$X_{\text{BuildRisk}}$	Valuation Office Agency estimate of property and social risk to a fire and rescue authority
$X_{\text{Primary2004}}$	Number of pupils between 5 and 11 years old
$X_{\text{ACORN risk}}$	Proportion of residents in houses with greater need of fire safety information
Pop	Registrar General projected population
X_{Area}	Area cost adjustments

Table 4. Key constants in formula (1)

Constant	*Definition*
4.4940954747	Smallest relative needs formula for each of the fire and rescue authorities divided by 10000
4.6998	1.6708 + 13/0.233: fire and rescue basic amount (reported as 1.6708), but in effect also contains the sum of the constants (13) added to each of the scaled deprivation indicators multiplied by the deprivation scaling value (0.233)
13	Total of terms added to ensure each deprivation indicator is positive
0.7788	Weighting for *per capita* coastline element
0.233	Weighting given to the sum of standardized 'risk' or 'deprivation' indicators
30669.4923	Weighting given to the number of *per capita* control of accidents and major hazards sites
865.2528	Weighting for property and social risk factors
0.4132	Weighting given to community fire safety

Finally, does the actual funding received depend largely on the subjective approval (or otherwise) of capital expenditure? Consider the rest of the formula:

$$Y_{\text{Debt}} = \frac{0.16188458389379}{\text{Pop2007}} \left\{ 0.054 \left(X_{\text{Debt}} + \frac{X_{\text{Exp}}}{2} - \frac{X_{\text{Debt}}}{50} \right) + \frac{X_{\text{Debt}}}{25} \right\} - 0.00000952457616 \qquad (2)$$

where X_{Debt} is debt and X_{Exp} is approved capital expenditure, and 0.054 is the assumed interest rate, 0.16188458389379 is a constant and 0.00000952457616 is the floor amount for debt handling.

Equations (1) and (2) are combined by

$$\text{RNF}_{\text{Fire}} = \frac{\text{Pop2009}}{1000000} \frac{Y_{\text{Fire}} + Y_{\text{Debt}}}{10000} \qquad (3)$$

Downloaded from https://academic.oup.com/jrsssa/article/174/3/547/7077760 by guest on 05 November 2025

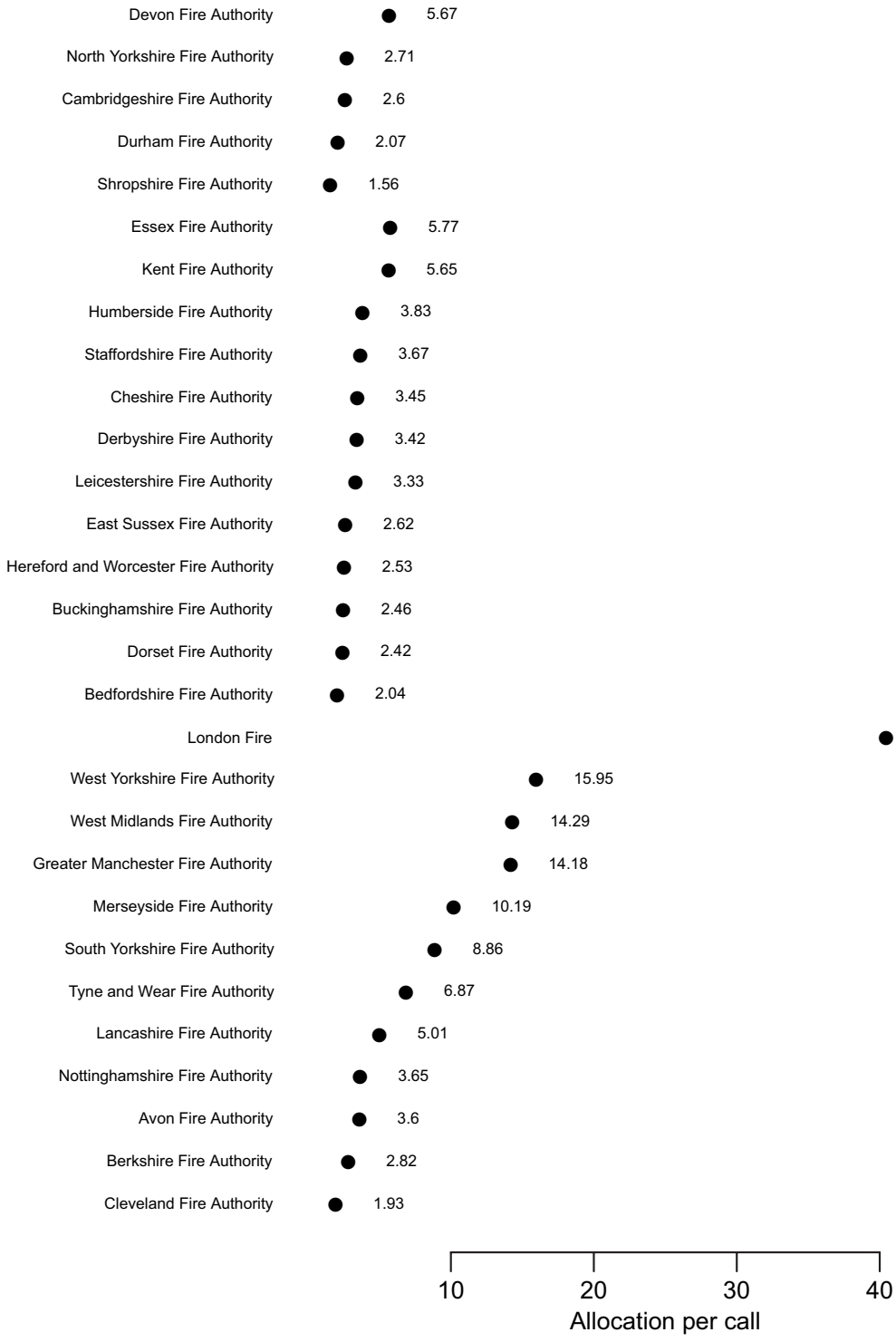

Fig. 3. Formula allocation per call received

Downloaded from https://academic.oup.com/jrsssa/article/174/3/547/7077760 by guest on 05 November 2025

to give a total relative needs formula score. Yet only large metropolitan areas have sufficient 'approved capital expenditure' to benefit from equation (2), and the reasons for this are not clear. In the overall 'four-block model that was reported by Gibson and Asthana, the Secretary of State can determine the proportion of money that is allocated directly *per capita* and the proportion that is allocated by the relative needs formula. How objective are these formulae and how much scope has the Secretary of State to influence funding?

The **authors** replied later, in writing, as follows.

Jane Galbraith and Mervyn Stone

We thank the organizers for combining our paper with another carrying the same message—the pressing need for government departments to change the way that they construct complex financial instruments. The discussants have given us a variety of valuable comments and most of the contributions severely stretch our ability to add further value. We thank them for their contributions and for the additional references provided.

Irini Moustaki has given all of us a useful overview and further suggests that latent variable modelling might be useful in developing better measures of need and supply. Similarly, Ian Hunt suggests that data mining, with or without a scientific basis, need not be an abuse of regression if used for exploratory, descriptive or even predictive purposes, provided that undue reliance is not placed on formal significance tests to validate the results. He agrees with us that the 'Combining age-related and additional needs' (CARAN) report (Morris *et al*., 2007) regression modelling is data mining done badly and we can agree with him that statistical modelling, e.g. using regression analysis, multilevel modelling or latent variable analysis as suggested by Irini Moustaki, can usefully uncover structure in a data set and provide background information. But, to echo Darrin Baines, there are no 'simple answers from complicated correlations'. Baines's interesting answers to the question 'why the health service still uses such a flawed method' provide insight into the interaction between policy makers and researchers.

Stephen Senn appears to assume that all would be well if only statisticians could agree on the solution. But it is, in the main, health economists who have been in the driving seats, supposing that their theory will convert the 'what "is" ' into the 'what "ought to be" '—a distinction that both he and Baines make.

Wynand van de Ven outlines an alternative system, the Dutch system with 'equalization payments' given to the insurer. Both he and Woody Caan throw valuable light on the proposed changes in the structure of the National Health Service. van de Ven highlights the possibilities that general practitioner consortia will try to commission from low cost, low quality, providers and to discriminate against high risk patients. Caan refers to the latter danger before emphasizing the need for general practitioner consortia to be geographically coherent and to employ statisticians: 'Amen' to that!

Colin Sanderson discusses some of the practicalities and difficulties in our tentative suggestion for research into direct measurement of need, concluding that, in the interim, 'adjusted utilization ... remains a practical approach'. We could agree with that if only it were openly acknowledged that the current 'adjustment' uses a more-or-less arbitrary combination of proxy variables and could be less equitable than, for example, starting with equal *per capita* allocation within each age group and only making such adjustments as could be shown to be justified.

We appreciate David Bartholomew's strong and well-informed support. His wide experience is at the back of the case that he makes for a more direct approach using a single index of need, presumably based on direct measures of morbidity as we have previously suggested. One of us is reminded of a year sharing an office with David (and may well have discussed this very issue with him) and finding a negative coefficient in the multiple-regression output of an IBM 1620 computer run—one that appeared to say that Irish farmers should pay other farmers to take some land off them.

Many of the variables that were used in the 'Allocation of resources to English areas' (AREA) (Sutton *et al*., 2002) and CARAN reports to measure need are far from what they appear to be at first sight. Nigel Williams reviews the housing domain index of deprivation and finds convincing evidence that it 'serves better as an identifier of exceptions than as a scaled measure of need'!

Like Nigel Williams, Greg Phillpotts is concerned about complexity and its possible unintended consequences. We agree that it is important to look at the calculated allocations but there are dangers. Firstly, the politicians may wish to tailor the allocations to favour individual or groups of primary care trusts (PCTs); secondly it is difficult to judge whether a set of outcomes is 'fair' without some criteria separate from those used in constructing it. However, Asthana and Gibson (2008) have found that the 2003–2006 target allocations might be unfair to rural and to more elderly populations, a matter on which Stone (2007)

Downloaded from https://academic.oup.com/jrsssa/article/174/3/547/7077760 by guest on 05 November 2025

found a degree of ambiguity in the Department of Health position as we explain in Appendix A of our paper.

We regret that none of the authors whose work we criticized contributed in person to the discussion. Instead, Hugh Gravelle, Stephen Morris, Nigel Rice, Peter C. Smith and Matthew Sutton submitted a short written contribution. They claim that the word limit prevents them from making a full rebuttal, but, by writing individually, the five of them could have had 2000 words! They say that our paper is 'replete with errors of fact, misinterpretations ... and misunderstandings ...'. But as they do not specify any errors and have declined our request for details it is diffcult to take this comment seriously.

However, we may have been misled by the term 'strategic health authority indicator' which we interpreted as denoting a dummy (0,1) variable analogous to the health authority dummy (0,1) variable in the AREA report. Hence, the details of our exposition involving strategic health authorities, e.g. in Section 1, page 519, and Section 4, pages 522–523, would apply to health authorities in the AREA report, but not to strategic health authorities in the CARAN report. This does not affect our basic argument that the regional indicators may measure need as well as supply, so fixing them at their means may remove information about need. It also does not weaken our technical *exposé* of the misleading inadequacy of the regression specification error test.

The third paragraph of the group's written contribution defends their work as explanatory, i.e. inferential—which may explain why they appear unconcerned that the inability of the latest formula (Morris *et al*., 2007) to include a term representing 'unmet need' (after 'exhaustive examination') was nevertheless the grounds for a Ministerial 'health inequalities' policy adjustment. When the Department of Health asked the CARAN research team to 'develop alternative future formulae that are suitable for allocations to PCTs' (Morris *et al*. (2007), page l3), it was not a request for inferential clues about the relationship between socio-economic variables and utilization (there is a plentiful supply of those) but for practical implementable formulae.

Although Gravelle and his colleagues admit that the CARAN team did not deal satisfactorily with the problem of 'unmet need', they do not acknowledge that the recommended regression equations do not deal satisfactorily with 'met need' either. A variable may have a large influence on the target formula while explaining little of the variation in recorded utilization. The relatively large residual variation about the fitted models may be due to 'met need', 'unmet need' or supply factors as well as to random variation.

We did not construct assumptions A and B as 'straw men' but as representations of what we took to be the conjectured theoretical framework on which the derived formulae were based. Indeed their fourth paragraph is confirmation that assumption B is needed to justify their recommendation of their 'predictive models' for calculating target allocations.

Gravelle and his colleagues wrongly claim that we said 'current knowledge cannot justify a departure from uniform *per capita* allocation'. What we said and still say is that any departure needs to be justified, especially when, as Joan Davis pointed out, the target allocation would give up to twice as much for each person in one PCT as in another. The issue is not whether differences in age profile or in deprivation justify differences in allocation but how large these differences should be. The conclusion of our paper was not that uniform per *capita* allocation should be the default, but that alternatives to the current formula need urgent investigation and that 'attention should be given to the machinery of government that contracts external expertise for research on departmental projects'.

We are particularly grateful to Joan Davis and Jon Spain for suggesting that the Royal Statistical Society is ideally placed to act as champions for 'ordinary people'. The Society is the only public body with 'both the national authority and the breadth of expertise' to scrutinize and provide independent advice on National Health Service funding formulae. We are, however, uncertain about Jon Spain's request that we 'refrain from being critical' on the pragmatic grounds that it 'just will not work'. Our hope is that pointing out the lack of justification for the allocation formulae in current use will spur researchers to investigate existing and new alternatives and the Department of Health to commission and use research more wisely.

Alex Gibson and Sheena Asthana

Our paper was originally submitted in late 2008 and was very much a product of its time. It refers primarily to the 2008–2009 settlement and focuses on issues that were then of particular concern to us. These were, first, the funding bias that flowed from the fact that below-threshold relative needs formula (RNF) values and tax base values were 'worth' 31% more that their above-threshold equivalents and, second, the hypersensitivity of the formula to changes in *per capita* RNF and tax base thresholds—the so-called 'Wokingham effect' which Stone in the discussion has now helpfully presented in concise algebraic form. We also hoped to convey something of the complexity and lack of transparency of the current approach,

Downloaded from https://academic.oup.com/jrsssa/article/174/3/547/7077760 by guest on 05 November 2025

and to argue that its serious methodological shortcomings were in no small part due to a lack of effective independent critical review.

There was, of course, much that we could not cover. Hewson and Halliday usefully demonstrate, for instance, that behind the complexity to which we refer lies yet more complexity. They examine the fire and rescue services formula in detail, highlighting *inter alia* an apparently erroneous use of one of the deprivation indicators. They could not, however, determine whether there was some unknown logic behind this and some other perplexing aspects of the formula because they were unable to obtain the necessary detailed technical exposition from the Department for Communities and Local Government. Such information should be readily available, and it is worth noting that very little background material is provided regarding the derivation of many of the other formulae that are used to determine RNF scores (Department for Communities and Local Government, 2011a).

Our principal concern, however, lay with the structure of the four-block model itself. Of particular significance are the need and resource equalization proportions. These are set by the Secretary of State without either explanation or justification although, as noted by Heiser, there are other areas 'where the statistical formulae are constrained by judgement'. Our contention, however, is that it is not sufficient simply to assert that these parameters are based on judgement. As described below, the need and resource equalization proportions have a profound effect on how resources are distributed between authorities and it is essential that their role, and the basis on which they are set, is made clear. If nothing else, silence provides no defence against the 'bad faith, improper motive or manifest absurdity' to which Heiser refers, whereas the outcomes that flow from the decisions made could easily raise just such accusations.

We explore elsewhere why the need and resource equalization proportions were set at 70.99% and −24.5% in 2006–2007, and at 73% and −26.6% in 2008–2009 (Gibson and Asthana, 2011). This analysis, in the absence of any explicit guidance or explanation from the Department for Communities and Local Government, was based on 'reverse engineering' the decision-making process. It is nevertheless clear that the four-block model was being manipulated to obtain predetermined funding outcomes. This, we argued, flies in the face of the basic tenets of formula funding in that, being outcome led, the approach cannot claim to offer an objective, rational or impartial approach to the distribution of the formula grant, It is not just, as Heiser puts it, that the model is *constrained* by judgement—it is completely *subservient* to the *post hoc* choice of the need and equalization proportions. This completely undermines the legitimacy of the local government funding formula, as can be demonstrated by examining the reasons for, and consequences of, the introduction of a new need equalization proportion as part of the recently announced 2011–2012 settlement.

In keeping with previous practice, the 2011–2012 settlement was preceded by a consultation paper describing the proposed changes to the formula funding methodology and requesting feedback from stakeholders. One of the issues that was raised was whether the need and equalization proportions should be set to ensure that the 'central allocation scaling factor' equals 1. The consultation paper was quite explicit about the importance of this:

> 'A change in the Central Allocation scaling factor away from one means that the amount below the threshold can have either more or less weight applied to it than the amount above the threshold. ... It has therefore been suggested that if the scaling factor for the Central Allocation is not very close to one, the distribution of grant will be distorted; and that this should be avoided' (Department for Communities and Local Government (2010), page 193).

Our paper strongly supports such concerns and, responding to the consultation, we agreed that a central allocation scaling factor of 1 was an essential (though not in itself sufficient) prerequisite for an equitable distribution of the formula grant. We were clearly not alone in this view in that, as reported in the Department's summary of representations,

> 'the vast majority of authorities that responded believe the scaling factor should be set close to one as this would increase the clarity, stability and fairness of the formula giving equal importance to both relative needs and resources above and below the minima' (Department for Communities and Local Government (2011b), page 14).

Anticipating such a response, the consultation paper proposed two options for obtaining a scaling factor of 1: first, by keeping the need equalization proportion at 73% and reducing the resource equalization proportion to −25.5%, or, second, by keeping the resource equalization proportion at −26.6% and raising the need equalization proportion to 74.6%. The Department's summary of representations notes that 'views were pretty evenly divided' between these two options before adding that

Downloaded from https://academic.oup.com/jrsssa/article/174/3/547/7077760 by guest on 05 November 2025

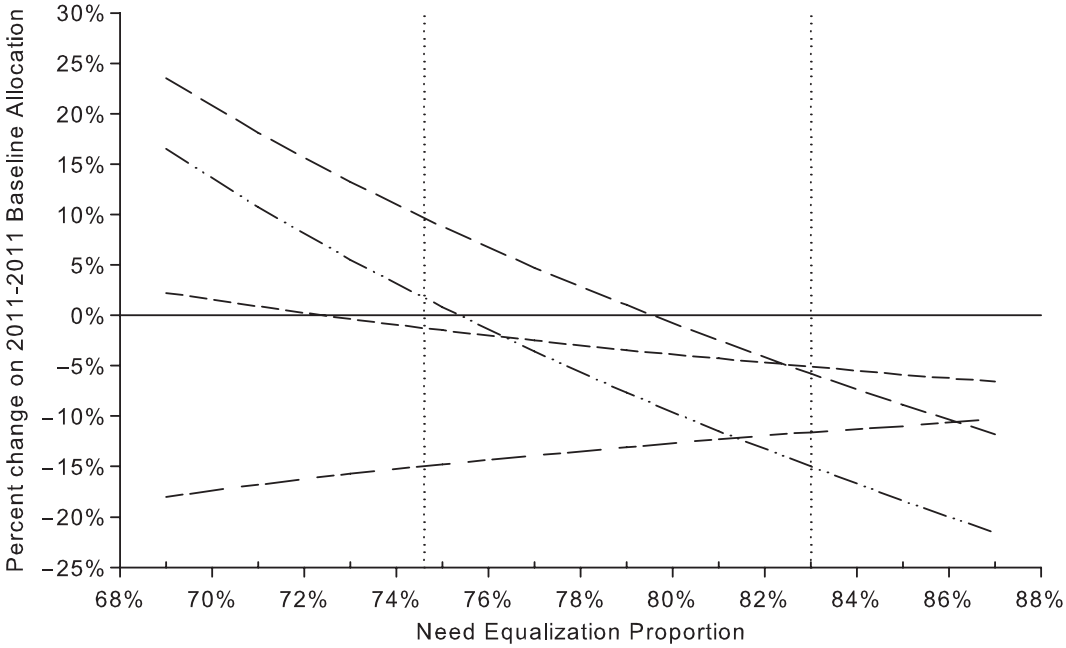

Fig. 4. Influence of the need equalization proportion on 2011–2012 floor group allocations: — – —, education and personal social services; – – –, fire; – ·· –, police; — —, shire districts

'whilst authorities support equalisation in principle, they would like to stress the importance of publishing the agreed principles behind any need or resource equalisation exercise'.

Yet when the 2011–2012 formula funding settlement was finally published it emerged that the Department for Communities and Local Government had decided to keep the resource equalization proportion unchanged at −26.6% but to increase the need equalization proportion to 83%. As this results in a scaling factor of 0.634—whereby below-threshold RNF and tax base scores are now worth only 63.4% of their above-threshold equivalents—the obvious question is why the department decided both to reject its own advice and to ignore the views of the 'vast majority of authorities'.

The answer, it seems, was that, in the face of a 9.9% reduction in the size of the formula grant, the original need equalization proportion of 74.6% would have resulted in what was, quite simply, an unacceptable funding outcome. At this level, as illustrated in Fig. 4, fire authorities and shire districts would have received respectively overall *increases* of 9.7% and 1.7% in their formula grant, whereas police authorities and education and personal social services authorities would have seen 1.3% and 15.0% *reductions* in their overall allocations. As these are aggregate figures, the Department for Communities and Local Government would have been in the unenviable position of having to explain why most authorities with responsibility for education and personal social services would have had to face double-digit reductions in funding at a time when many shire districts, and almost all fire authorities, would have been in receipt of double-digit increases in allocations.

Fig. 4 illustrates the quandary that the department faced. A need equalization proportion had to be found which resulted in a distribution between the four floor groups that was commensurate with the overall reduction in funding and, for whatever detailed reason, a value of 83% was chosen. The implications for the various floor groups can be read from Fig. 4 and it is clear that, in effect, allocations to police, fire and shire district authorities were sacrificed to ensure that they were not out of line with what could be directed towards the education and personal social services authorities.

Three points are worth emphasizing. First, these figures refer to post-damping allocations rather than predamping targets. They concern, in other words, 'real' money. Second, the link between the need equalization proportion and the resource flowing to each of the floor groups is a direct and inevitable consequence of the underlying pattern of authority level RNF and tax base values, allied with the fact that the floor

Downloaded from https://academic.oup.com/jrsssa/article/174/3/547/7077760 by guest on 05 November 2025

Table 5. Winners and losers when the need equalization proportion is set at 83% rather than 74.6%

Local authority	*Floor group*†	*Formula grant (£ million) when need equalization proportion is*		*Change in allocation*	
				Absolute (£ million)	*Relative (%)*
		74.6%	*83%*		
(a) Most advantageous absolute reallocations					
Birmingham	EDUPSS	663.92	691.24	27.32	4.1
Kent	EDUPSS	300.93	316.14	15.21	5.1
Lancashire	EDUPSS	319.97	333.82	13.85	4.3
(b) Least advantageous absolute reallocations					
Greater London Authority—police	POL	2071.30	2025.58	−45.72	−2.2
Greater London Authority—fire	FIR	278.17	251.64	−26.54	−9.5
West Midlands Police	POL	522.12	497.78	−24.33	−4.7
(c) Most advantageous relative reallocations					
Enfield	EDUPSS	123.10	131.01	7.90	6.4
Redbridge	EDUPSS	95.12	101.06	5.94	6.2
Dorset	EDUPSS	61.84	65.54	3.70	6.0
(d) Least advantageous relative reallocations					
East Dorset	SD	3.78	2.60	−1.18	−31.2
Christchurch	SD	2.75	1.92	−0.83	−30.2
Hart	SD	4.13	2.95	−1.18	−28.6

†EDUPSS, education and personal social services; POL, police; FIR, fire; SD, shire districts.

damping block was set up to prevent cross-subsidization between the floor groups. And, third, the same service needs (and same capacity to raise resource) exist whatever the funding scenario. The choice of the need equalization proportion thus has a direct bearing on the ability of individual authorities to meet the needs of the populations that they serve.

It is worth, therefore, considering how individual authorities fared in the face of this decision to abandon the proposed 74.6% need equalization proportion in favour of one set at 83%. This is complicated by the fact that the precise effect on individual authorities depends on each floor group's funding guarantee, and these are also set 'by judgement'. They can never be more advantageous than the floor group's aggregate change in funding (at which point all authorities within the group would receive precisely the same change in funding), but less advantageous guarantees are possible—in which case there is greater diversity in funding outcomes as authorities with target allocations above the guarantee make up a proportion of the difference between the guarantee and the target. This proportion is known as the 'floor group scaling factor' and, to isolate the effect of changing the need equalization proportion, we can compare allocations while applying the floor group scaling factors that were used in the 2011–2012 settlement.

Table 5 thus lists those authorities with the highest absolute and relative gains and losses when the need equalization proportion is set at 83% rather than 74.6% Overall, education and personal social services authorities gain £730 million (3.9%) whereas police, fire and shire district authorities lose respectively £320 million (−3.9%), £164 million (−14.1%) and £246 million (−16.5%).

The Department for Communities and Local Government's manipulation of the 2011–2012 funding settlement demonstrates the four-block model's fundamental weakness; that it cannot be used to 'equalize' the capacity of local authorities to meet the service needs of their populations. Does, for instance, the London Borough of Enfield require 33 times as much revenue support as East Dorset District Council (as appears necessary when the need equalization proportion is set at 74.6%), or 50 times as much (which is what emerges when the need equalization proportion is set at 83%)? The four-block model provides no way of knowing, and what actually emerges is an incidental consequence of the Department for Communities and Local Government's search for an overall funding outcome that makes it *appear* that the four floor groups are being treated equitably.

Downloaded from https://academic.oup.com/jrsssa/article/174/3/547/7077760 by guest on 05 November 2025

The four-block model is anathema to the idea of *formula* funding because it dispenses with the idea of an objective process, and the fact of the matter is that the Department for Communities and Local Government may as well have allocated funds without any recourse to the measurement of local service needs and capacity to raise income. The four-block model needs to be replaced as a matter of urgency, and it must certainly not be used to inform the development of any new mechanism for funding the devolved governments and assemblies of the UK (Cuthbert and Cuthbert). As for Senn's reproach that we have left unsaid what we propose should be done, that would, unfortunately, require far more space than is available to us here—though we are satisfied that it is indeed possible to construct an equitable, simple and transparent formula which would avoid the manifest absurdity which permeates the current local government funding formula and, as Heiser hopes, be more comprehensible to key officers and members.

References in the discussion

Ashton J. R. (1978) R.A.W.P. and the achievement of equity in the national health service. *Publ. Hlth*, **92**, 190–196.

Asthana, S. and Gibson, A. (2008) Healthcare equity, health equity and resource allocation: towards a normative approach to achieving the core priniciples of the NHS. *Rad. Statist.*, no. 96, 6–26.

Baines, D. L. and Parry, D. J. (2000) Analysis of the ability of the new needs adjustments formula to improve the setting of weighted capitation prescribing budgets in English general practice. *Br. Med. J.*, **320**, 288–290.

Beaulieu, N., Cutler, D. M., Ho, K., Ishma, G., Lindquist, T., Nelson, A. and O'Connor, P. (2006) The business case for diabetes management for managed care organizations. *Form Hlth Econ. Poly*, **9**, 1–36.

Bennett, R. J. and Tan, K. C. (1979) Allocation of the Rate Support Grant by the use of methods of optimal control. *Environ. Planng*, **11**, 1011–1027.

Bevan, R. and van de Ven, W. P. M. M. (2010) Choice of providers and mutual healthcare purchasers: can the English National Health Service learn from the Dutch reforms? *Hlth Econ. Poly Law*, **5**, 343–363.

Boyle, D. (2001) *The Tyranny of Numbers: Why Counting Can't Make Us Happy*. London: Harper Collins–Flamingo.

Caan, W. (2006) "Be healthy!" if every child matters. (Available from `www.BMJ.com`.)

Caan, W. (2009) Learning disability. *Br. J. Gen. Pract.*, **59**, 209.

Cook, W. R. (1976) How the University Grants Committee determines allocations of recurrent grants. *J. R. Statist. Soc.* A, **139**, 374–384.

Department for Communities and Local Government (2008) Statistical release, updated Council Tax levels set by Local Authorities in England—2008-09, p. 13. Department for Communities and Local Government, London. (Available from `http://www.communities.gov.uk/documents/statistics/pdf/1289074.pdf`.)

Department for Communities and Local Government (2010) *Local Government Finance Formula Grant Distribution: Consultation Paper*. London: Stationery Office. (Available from `http://www.local.communities.gov.uk/finance/1112/sumcon/index.htm`.)

Department for Communities and Local Government (2011a) *Local Government Finance Report (England) 2011/2012*. London: Stationery Office. (Available from `http://www.local.communities.gov.uk/finance/1112/lgfr1112s/index.htm`.)

Department for Communities and Local Government (2011b) Consultation on Formula Grant distribution: summary or representations. Department for Communities and Local Government, London. (Available from `http://www.local.communities.gov.uk/finance/1112/sumcon/conrepanalysis.pdf`.)

Department of Environment, Transport and the Regions (2000) *Indices of Deprivation 2000*, pp. 24, 25, 42. London: Department of Environment, Transport and the Regions. (Available from `http://www.communities.gov.uk/documents/regeneration/pdf/131306.pdf`.)

Department of Health (1989) *Working for Patients*. London: Her Majesty's Stationery Office.

Department of Health (2008) Exposition book 2009-10 and 2010-11, Table 3. Department of Health, London. (Available from `http://www.dh.gov.uk/en/Managingyourorganisation/Financeandplanning/Allocations/DH_091850`.)

Fillingham, D. (2010) NHS resource allocation formula. ACRA, Salford. (Available from `http://www.dh.gov.uk/en/Publicationsandstatistics/Lettersandcirculars/Dearcolleagueletters/DH_122670`.)

Gibson, A. and Asthana, S. (2011) A tangled web: complexity and inequality in the English Local Government Finance Settlement. *Local Govern. Stud.*, to be published.

Hand, D. J. (2008) *Statistics*. Oxford: Oxford University Press.

Hendry, D. F. and Granger, C. W. J. (2005) A dialogue concerning a new instrument for econometric modelling. *Econometr. Theor.*, **21**, 278–297.

Her Majesty's Government (2010) Equity and excellence: liberating the NHS. *White Paper*.

Kendall, M. (1953) The analysis of economic time series. *J. R. Statist. Soc.* A, **116**, 11–34.

Leamer, E. E. (1983) Let's take the con out of econometrics. *Am. Econ. Rev.*, **73**, 31–43.

Lovell, M. C. (1983) Data mining. *Rev. Econ. Statist.*, **65**, 1–12.

Downloaded from https://academic.oup.com/jrsssa/article/174/3/547/7077760 by guest on 05 November 2025

McCarthy, M. (1986) A fairer deal from RAWP. *Br. Med. J.*, **292**, 289–290.
McCloskey, D. N. (1985) The loss function has been mislaid: the rhetoric of significance tests. *Am. Econ. Rev.*, **75**, 201–205.
McCloskey, D. N. (2008) *The Cult of Statistical Significance*. Ann Arbor: University of Michigan Press.
Morris, S., Carr-Hill, R., Dixon, P., Law, M., Rice, N., Sutton, M. and Vallejo-Torres, L. (2007) Combining age-related and additional needs. *Report RARP 30*. Department of Health, London.
Noble, M., Wright, G., Dibben, C., Smith, G. A. N., McLennan, D., Anttila, C., Barnes, H., Mokhtar, C., Noble, S., Avenell, D., Gardner, J., Covizzi, I. and Lloyd, M. (2004) *Indices of Deprivation 2004*, p. 151. London: Office of the Deputy Prime Minister. (Available from `http://www.communities.gov.uk/documents/communities/pdf/131209.pdf`.)
Prinsze, F. J. and van Vliet, R. C. J. A. (2007) Health-based risk adjustment: improving the pharmacy-based cost group model by adding diagnostic cost groups. *Inquiry*, **44**, 469–480.
Rozewicz, L., Caan, W. and Johns, A. (1992) Managing drug misuse in general practice. *Br. Med. J.*, **304**, 1442.
Sanderson, C., Hunter, D., McKee, M. and Black, N. (1997) Limitations of epidemiologically based needs assessment: the case of prostatectomy. *Med. Care*, **35**, 669–685.
Senn, S. J. and Shaw, H. (1978) Resource allocation: some problems in applying the national formula to area and distict revenue allocations. *J. Epidem. Commty Hlth*, **32**, 22–27.
Spanos, A. (1999) *Probability Theory and Statistical Inference*. New York: Cambridge University Press.
Spanos, A. (2000) Revisiting data mining: 'hunting' with or without a license. *J. Econ. Methodol.*, **7**, 231–264.
Stam P. J. A. and van de Ven, W. P. M. M. (2007) Evaluatie risicoverevening: prikkels tot risicoselectie? *Report*. Erasmus University Rotterdam.
Stam P. J. A. and van de Ven, W. P. M. M. (2008) De harde kern in de risicoverevening. *Econ. Statische Ber.*, Feb. 22nd.
Stone, M. (2007) In *NHS Deficits, First Report of Session 2006-07*, vol. II, supplementary evidence, Ev 168–172. London: Stationery Office.
Sutton, M., Gravelle, H., Morris, S., Leyland, A., Windmeijer, F., Dibden, C. and Muirhead, M. (2002) Allocation of resources to English areas: individual and small area determinants of morbidity and use of healthcare resources. *Report RARP 26*. Department of Health, London.
Timmins, N. (2010) Letter from Britain: across the pond giant new waves of health reform. *Hlth Aff.*, **29**, 2138–2141.
van de Ven, W. P. M. M. and Ellis, R. P. (2000) Risk adjustment in competitive health insurance markets. In *Handbook of Health Economics* (eds A. J. Culyer and J. P. Newhouse), ch. 14, pp. 755–845. Amsterdam: Elsevier.
van de Ven, W. P. M. M., van Vliet, R. C. J. A. and Lamers, L. M. (2004) Health-adjusted premium subsidies in the Netherlands. *Hlth Aff.*, **23**, 45–55.
Willcock, G. K. (1979) The prevalence of osteoarthrosis of the hip requiring total hip replacement. *Int. J. Epidem.*, **8**, 247–250.

Downloaded from https://academic.oup.com/jrsssa/article/174/3/547/7077760 by guest on 05 November 2025

View publication stats